DOCUMENT RESUME

ED 197 323 CS 005 860

AUTHOR Thompson, Murray D.; Block, Karen K.
TITLE The Effects of Three Types of Practice Formats and Two Degrees of Learning on the Spelling Performance of Elementary School Students.
INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.; National Science Foundation, Washington, D.C.
REPORT NO LRDC-1979/19
PUB DATE 79
GRANT NSF-GJ-540X-1
NOTE 17p.; Small print may be marginally legible.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *Educational Research; Grade 5; Grade 6; Intermediate Grades; Reading Research; *Recall (Psychology); *Spelling; *Spelling Instruction; Teaching Methods; Visual Learning; *Word Recognition

ABSTRACT

A study was conducted that used three formats for spelling practice, two multiple-choice or recognition conditions based on the Simon model and one production or recall condition. (The Simon generate-and-test model of spelling suggests that spellings can be learned from reading or visual exposure to the correct spelling.) The multiple-choice formats contained distractors that were either plausible graphemic variations of the target word or misspellings of different words. Fifth and sixth grade students learned 15 words whose spellings they had missed on a pretest in a small group instructional setting in which words had been practiced for either three or six trials. Students were blocked into high and low ability groups and the effects of variations in training, degree of learning, and ability were assessed on an immediate and delayed generate-and-test measure and on a delayed recognition test. Results showed that all training groups learned more spelling than did the control group. However, the recall condition was superior to each recognition condition on each test. High ability students outperformed low ability students on all three tests. (Author/MKM)

THE EFFECTS OF THREE TYPES OF PRACTICE FORMATS AND TWO DEGREES OF LEARNING ON THE SPELLING PERFORMANCE OF ELEMENTARY SCHOOL STUDENTS

Murray D. Thompson and Karen K. Block

Learning Research and Development Center

University of Pittsburgh

1979

This research was supported by a grant from the National Science Foundation (NSF-GJ-540X-1) and by the Learning Research and Development Center, supported in part as a research and development center by funds from the National Institute of Education, U.S. Department of Health, Education, and Welfare. The opinions expressed in this article do not necessarily reflect the position or policy of the sponsoring agencies and no official endorsement should be inferred. The authors wish to thank Mr. Arthur E. Fleeger and the teachers of Sickman Elementary School, Baldwin-Whitehall School District, for their extensive cooperation. Thanks is also due Carol E. Wildemann of the University of Pittsburgh Office of Measurement and Evaluation for help in data analysis and to John Levine for helpful editorial comments. Portions of these data were presented at a Meeting of the American Educational Research Association in April 1975.

THE EFFECTS OF THREE TYPES OF PRACTICE FORMATS AND TWO DEGREES OF LEARNING ON THE SPELLING PERFORMANCE OF ELEMENTARY SCHOOL STUDENTS

Murray D. Thompson and Karen K. Block

Learning Research and Development Center
University of Pittsburgh

According to the Simon and Simon (1973) theory of spelling performance, some words (those that are partially learned) are spelled through the use of visual recognition information that has been acquired through reading. Spelling is considered to be a generate-and-test process in which knowledge of phoneme-grapheme correspondences is used to generate trial spellings which are written then tested against the stored word recognition information that is retrieved upon viewing an approximate spelling (Simon & Simon, 1973, p. 130). The major empirical support for the generate-and-test account of spelling comes from the serial position effect found for spelling errors (Kooi, Schutz, & Baker, 1965). The notion is that since recognition information is most clear at the start and end of words, and fuzzy in the middle, incorrect spellings will pass the test when errors exist in the middle positions rather than at the start or end. Therefore, the majority of spelling errors should be in the middle of the word, hence the serial position effect.

There are two difficulties with the above argument. First, the serial position effect may be a phenomenon of recall rather than recognition. That is, it may simply be due to heightened recall of the start and ending of words. When generating a spelling of a word they have seen before, students may be able to recall (and subsequently write) the start and ends of it easily and correctly while having difficulty

remembering the middle. This latter suggestion does not require the positing of a post-generation recognition "test" phase. The second difficulty with the argument is that the spelling error data cited in support of the model were not collected under conditions where there were assurances that proofreading was a regular occurrence. A requirement of the model (see p. 130) is that the test phase must be initiated by presentation of the spelling of a word; however, it is unreasonable to assume students subject each spelling they have generated to a "test." Indeed, classroom spelling teachers must regularly remind their students to check the spellings they have generated.

The distinction between the above two accounts is important both for understanding spelling processes and also because the two accounts seem to have quite different implications for instruction in spelling. The Simons' (1973) suggest that much of spelling can be learned by reading because by reading the words one needs to spell, one builds up recognition information, thereby facilitating spelling performance. They argue for the importance of correlated vocabularies in reading and spelling. This in turn implies that standard forms of teaching spelling by requiring production and giving feedback, derived from viewing spelling as fundamentally a recall rather than recognition process, may be less useful than we otherwise might think. The Simons' suggestion that spelling competence can be strongly elevated through "reading" (i.e., visual exposure to correct spellings of words) is in need of empirical test. This study provides data to assess the relative amount of achievement that results from multiple-choice (recognition) training and constructed response (recall) training.

It was hypothesized that if practice in recognizing words resulted in greater spelling achievement, then recognition must be considered a fundamental spelling process. By contrast, if practice in recalling words resulted in superior achievement, then spelling must be considered fundamentally a recall rather than recognition phenomenon. A

multiple-choice format was used to provide practice in "recognizing" the correct spelling; a constructed response format was used to provide practice in "recalling" the correct spelling. We refer to the two types of training procedures as "recognition" and "recall" because the tasks involved are identical to the traditional methods for testing recognition and recall. Our names for the training conditions are based on external conditions for learning rather than internal ones. Each training format may elicit in some students both recognition and recall processes because the training tasks permit only limited control over spelling processes.

Two recognition training procedures were developed on the basis of hypotheses suggested by Simon and Simon (1973). They suggested that spellers can increase spelling proficiency by developing more elaborated templates, through visual exposure to the words. Two forms of recognition training provided such exposure while varying in the nature of the distractors used in multiple-choice tasks. In the high recognition condition, the distractors were misspellings of the target word, many of which were the most frequent misspelling committed on the pretest or other misspellings formed by substituting alternative plausible graphemes for selected phonemes in the word. By contrast, the low recognition condition used misspellings of completely different words as distractors. Thus, both recognition training conditions provided the requisite visual spelling information, but differed in the nature of misspelled distractors. It was predicted that high recognition training would produce better spelling performance than low recognition training. This prediction was based on the assumption that requiring a choice among plausible alternative spellings would facilitate the learning of the correct spelling, perhaps by forcing attention to relevant orthographic features.

The third practice condition (recall) was a standard method of spelling practice. Students attempted to spell a word and were then

shown the correct spelling. It was hypothesized that if recognition practice was in fact a superior form of spelling practice, more spellings should be acquired in both recognition conditions than in the recall condition.

In addition to variations in training procedure, the effects of two other variables were assessed in this study. Training was carried to three or six trials, hence degree of learning was varied. It was expected that six practice trials would produce more learning than three. Students were blocked into high and low ability spellers to assess any interactions that ability might have with treatment.

Method

Subjects

The subjects were 135 fifth- and sixth-grade students from a suburban elementary school. They were chosen from a larger population of 165 students on the basis of pretest performance and the spelling subtest of the Stanford Achievement Test.

Students were eligible for the study if they incorrectly spelled at least 85% of the words on a 60-item pretest and misspelled all 15 words that were used as learning items in the study.

The median score (obtained by students qualifying for the experiment) on the spelling subtest of the Stanford Achievement Test was used to identify high and low ability spellers. The range of scores in grade equivalents was 3.8 to 7.2 and the median was 5.7. Those above the median were considered high ability spellers and, inversely, those below the median were considered low ability spellers.

Design

An equal number of fifth and sixth graders were used. Subjects were randomly allocated to the experimental conditions such that an

equal number of fifth-grade, sixth-grade, high ability, and low ability spellers were in each training condition. Three training conditions were used in the study: high recognition, low recognition, and recall. Degree of learning (three versus six trials) varied factorially with training. Subjects were assigned to groups so that five low and five high ability, and five fifth and five sixth graders appeared in each combination of training conditions with degree of learning condition. There was a total of 120 students included in the main design which had four between factors (grade, ability, training, degree of learning) and one within factor (tests). A control group that received no training was also run; it consisted of seven low ability and eight high ability fifth and sixth graders.

Apparatus

An opaque projector and standard movie screen were used to present stimulus items and feedback.

Word Materials

Sixty spelling words difficult for sixth graders were selected from the Iowa Spelling Scale (Greene, 1954) and used in a pretest. The 15 most frequently missed spellings in the pretest were used as the learning items for the study.

Multiple-choice items with four distractors were developed for each of the 15 learning items for the low recognition and high recognition conditions. In the high recognition condition, four highly similar spellings of the target words were used, only one of which was the correct spelling. The three distractor items for a target word were the most frequent misspelling that occurred on the pretest and two graphemic variants (e.g., forein, foriegn, forien for foreign). Similarly, the low recognition condition had four spellings from which to

choose. However, the distractor items were misspellings of completely different words. The distractors varied in number of pronounceable chunks and no single pronounceable chunk found in the target word was contained in any distractor. An example set of choices for foreign was: restrant (2 chunks), agreculture (3 chunks), and hijean (2 chunks).

Procedure

A dictated pretest consisting of 60 spelling words from the New Iowa Spelling Scale (Greene, 1954) was administered to all fifth- and sixth-grade students. Training began seven days after pretesting. All training was conducted in a classroom setting. Twenty students were exposed to training simultaneously. These students were assigned to the same training and degree of learning conditions, but were from different grade levels. Hence, 10 fifth and 10 sixth graders were tested in each session. Two students, who did not qualify for the experiment, aided the experimenter. These students were used as runners to collect subjects for training and return them to their respective classrooms at the completion of training. Through the cooperation of the classroom teachers, students were not exposed to spelling instruction, other than that associated with experimental training, for the three-week duration of the experiment.

In the recognition conditions, the experimenter projected a multiple-choice array of four spellings on a screen, pronounced the target word, used it in a sentence, and pronounced it again. Students wrote the letter (a, b, c, d) of the spelling they thought correct. The experimenter then showed the correct letter and spelling of the word. Students checked their answers marking a "c" for correct or a "/" for incorrect. Training continued in this way for all 15 target words. The list was recycled in the same fashion for either three or six trials. The order of word presentation and the position of the correct spelling and distractors for

each item was randomized on each presentation. The conduct of recall training was similar, except students wrote a complete spelling then were shown the correct spelling and had to check it "c" or "/". Degree of learning (three or six trials) was also varied in recall. Students in the control group received no training, only the tests.

Six classroom training sessions were conducted in sequence in the same room on the same day. At the end of each training session, that is, at the end of three or six times through the list, a "generate-and-test" posttest was administered to each training group. Each of the 15 words was, as above, pronounced and used in a sentence, and pronounced again. The subjects had the option to write either one or two spellings for each item. If two spellings were written, the subjects underlined the spelling they thought more correct (which actually could be correct or not correct). Ten days later, two retention tests were given to all groups, including the control group, on a single day. One test followed the generate-and-test format used with the immediate posttest. The other test was a multiple-choice test. Distractors on the multiple-choice test were the most frequent misspelling, a misspelling of a completely different word, and a graphemic variant of the target word. Order of administration of the retention tests was counterbalanced. No feedback was given until both tests were completed.

Results and Discussion

Control versus Training

To determine whether the various training procedures produced learning of the words, the training groups were compared to the control group. Scores on the retention test were selected for analysis because retention is the practically significant measure of the effects of training. A one-way analysis of variance computed on the number of correct spellings produced in the training and control conditions on the generate-

and-test retention test showed a significant main effect (F = 15.97; df = 3,131; $p < .01$). The mean number correct was 6.73 in the recall condition; 3.21 in high recognition; 3.10 in low recognition; and .33 in the control group. A Scheffé post hoc analysis showed all three training groups produced significantly more ($p < .01$) correct spellings than the control. Clearly, all three training conditions produced learning of spelling words. This means that spellings are indeed learned from "reading-like" situations.

Analysis of Accuracy

A second analysis determined the effects of the variables in the main design. The effects of training, degree of learning, and ability were assessed through an analysis of variance on the number of correct spellings produced on each of the three tests, immediate and delayed generate-and-test tests and the delayed multiple-choice test. On the generate-and-test tests, students wrote one or two spellings for a word, then underlined the one they thought correct. The underlined spelling was scored correct or incorrect. Total correct was computed by adding correct single spellings and correct underlined spellings. Number correct on the multiple-choice test was the number of correctly selected spellings. The results of a repeated measures analysis of variance showed significant main effects for ability (F = 69.49; df = 1,108; $p < .001$), training (F = 37.30; df = 2,108; $p < .001$), tests (F = 256.19; df = 2,216; $p < .001$), and one significant interaction between training and tests (F = 18.48; df = 4,216; $p < .001$). High ability spellers (mean total correct over all tests = 8.1) outperformed low ability spellers (mean total correct over all tests = 4.3). Degree of learning had no effect taken alone or in combination with the other variables. Six "massed" trials had the same effect as three "massed" training trials in producing correct spellings.

Table 1 displays the means for each combination of training and test conditions.

Table 1

Mean Number Correct for Each Training and Test Condition

	Training			
Tests	High Recognition	Low Recognition	Recall	Total
Generate-and-test posttest	4.00	2.38	9.30	5.22
Generate-and-test retention test	3.28	3.10	6.73	4.37
Recognition retention test	7.82	8.45	10.90	9.06
Total	5.02	4.64	7.97	

As can be seen from Table 1, scores were higher on the recognition retention test relative to the generate-and-test tests and the recall condition produced the most learning relative to the other two training conditions. The locus of the interaction effect was determined by a simple effects analysis and Scheffé post hoc comparisons. A simple effects analysis showed there were significant differences between all three training conditions at each level of test ($F = 64.68$, $df = 2, 45$ for the generate-and-test posttest; $F = 20.61$, $df = 2, 45$ for the generate-and-test retention test; $F = 13.03$, $df = 2, 45$ for the multiple-choice test; all $p < .001$). Significant effects were also found among the tests at each level of training condition ($F = 81.79$, $df = 2, 216$ for high recognition; $F = 150.65$, $df = 2, 216$ for low recognition; and $F = 60.72$, $df = 2, 216$ for recall; all $p < .001$). Scheffé comparisons between the number correct in the three training conditions for each test showed recall training produced significantly more ($p < .01$) correct spellings than

high, or low, recognition training. The two recognition training conditions did not differ on any test. The superiority of constructed response training over multiple-choice training therefore held for all three tests. The Scheffé procedure also revealed performance in each of the three training conditions was the highest on the multiple-choice test, significantly higher ($p < .01$) than on either generate-and-test. This result is consistent with standard findings of superiority for recognition over recall (see also Hollingworth's [1918] stage model). The locus of the test effect within training conditions was a significant loss of information from immediate to delayed generate-and-test tests in the recall condition, while no significant loss was observed for either recognition condition. This could be a "bottoming out" effect since the amount initially acquired in the recognition conditions was relatively small in the first place.

Multiple versus Single Production

Another analysis was done to determine whether variation in training conditions influenced the degree to which multiple as contrasted to single spellings were produced on the tests. This could be construed as a rough index of the amount of effort necessary to produce the spellings requested on the tests to satisfy personal levels of confidence in a response. Table 2 shows for each training condition the words for which alternative spellings were generated out of a possible maximum of 15. The means were calculated by adding the number of words each student wrote multiple spellings for, then dividing by the total number of students.

It is clear from the means in Table 2 that multiple spellings were not very frequent, except perhaps in the low recognition condition on the delayed test when an average of about three words required multiple spellings. Inspection of the means in Table 2 shows that in the case of

most practical importance, the generate-and-test retention test, the low recognition condition required the most effort, more than high recognition and recall, when effort is measured in terms of frequency of multiple spellings.

Table 2

Mean Number of Alternatives and Standard Deviation as a Function of Practice Condition and Tests Administered

Tests Administered	High Recognition		Low Recognition		Recall	
	Mean	SD	Mean	SD	Mean	SD
Generate-and-test posttest	1.78	1.6	.4	.9	.78	1.0
Generate-and-test retention test	1.50	1.6	3.32	2.6	1.9	2.0

One might also be concerned with the relative effectiveness of the treatments had only single spellings been permitted on the tests (the usual condition for spelling tests) and with the amount of gain (or loss) contributed by the opportunity to do multiple spellings.

Table 3 shows the mean correct for each training condition and each test when (a) all words are considered and the underlined spelling scored correct or incorrect, or (b) only the words having single spellings are considered. Clearly, the pattern of superiority among training conditions is the same when all spellings are considered (total) or when only those without alternatives are considered. Furthermore, the means correct with, or without, alternatives are very close, indicating that the opportunity to do multiple spellings contributed

little to the overall mean correct. At least as far as group means are concerned, the chance to do multiple spellings contributed little to a gain in correct spellings. T tests, comparing mean correct with and without alternatives revealed no significant gains, nor are there practical gains due to the use of alternatives.

Table 3

Mean Number of Correct Spellings and Standard Deviations When Correct Spellings That Used Multiple Alternatives Are Removed and the Means Are Calculated Without Alternatives

		Test Administered			
		Posttest		Retention Test	
Training Conditions	Spelling Correct	Mean	SD	Mean	SD
Low recognition	Total	2.38	1.9	3.10	2.0
	Without recognition	2.35	1.9	2.75	2.0
High recognition	Total	4.00	2.6	3.28	2.6
	Without alternatives	3.72	2.6	3.00	2.7
Recall	Total	9.30	3.3	6.73	3.8
	Without alternatives	8.92	3.1	6.12	3.4

General Discussion

The data of the present study showed that spelling can indeed be learned from "reading-like" situations, but subsequent achievement is inferior to constructed response training. The assertion of the Simon model that spellings of words can be learned through visual exposure to the correct spelling is confirmed by these data. However, there is a clear advantage to practice in producing the correct spelling relative to simply selecting it from an array. Thus, while

recognition plays a role in spelling, it is best viewed as a secondary and/or supplementary process rather than a primary process. If the learning objective is to increase a child's spelling vocabulary, then drills using constructed responses are effective procedures to use.

Our data also showed no difference in amount of achievement between high and low recognition conditions. This finding seems to imply that the set of distractors did not function in the manner intended, that is, the "close in" distractors did not force more attention to the orthographic composition of the word, which in turn would presumably lead to greater achievement. It would seem, therefore, that the composition of the set of distractors would be unimportant, as long as they consisted of other misspelled words. Our feeling, however, is that such a conclusion is unwarranted because the data revealed a possible disadvantage to the use of "far out" distractors. The mean number of alternative spellings generated on the generate-and-test retention test tended to be greater in the low recognition condition. If this finding were replicated in a situation in which children were accustomed to generating alternative spellings on a regular basis, as a part of spelling procedure, it could possibly indicate that retrieval of the recognition template must be strongly cued by the presence of more than one alternative spelling. Thus, students in the low recognition condition are less confident in recognizing the correct production. Such a suggestion ought to be pursued in future research.

Finally, our data did not show evidence of a practice effect, for three "massed" trials were equivalent in amount of learning produced to six "massed" trials. Possibly, students' motivation for learning had decreased by the later trials, thereby decreasing the amount of learning acquired in the later trials. A graph of the course of acquisition collapsing across treatment groups showed that the number correct steadily increased from trials 1 to 3 with an asymptote reached

at trail 4. Thus, three trials are the limit for deriving benefits from massed practice. The data underscore the importance of distributed practice in spelling and when more than three trials are to be spent learning spelling lists, spacing should begin after trial 3.

References

Greene, H. A. The New Iowa Spelling Scale. Iowa City: State University of Iowa, 1954.

Hollingworth, L. S. The psychological examination of poor spellers. Teachers College Record, 1918, 126-132.

Kooi, B. Y., Schutz, R. D., & Baker, R. L. Spelling errors and the serial position effect. Journal of Educational Psychology, 1965, 56, 344-346.

Simon, D. P., & Simon, H. A. Alternative uses of phonemic information in spelling. Review of Educational Research, 1973, 43, 115-137.